RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/019,282
Source: PCT/10
Date Processed by STIC: 10/19/04

ENTERED

PCT

RAW SEQUENCE LISTING DATE: 10/19/2004
PATENT APPLICATION: **US/10/019,282** TIME: 17:29:38

Input Set : **D:\217770US0PCT.txt**
Output Set: **N:\CRF4\10192004\J019282.raw**

```
  3 <110> APPLICANT: SHIBATA, TAKASHI
  4       NOGUCHI, YUJI
  5       YAMASHITA, MICHIO
  7 <120> TITLE OF INVENTION: GENE ENCODING CYCLIC LIPOPEPTIDE ACYLASE AND EXPRESSION OF
THE
  8       SAME
 10 <130> FILE REFERENCE: 217770US0PCT
 12 <140> CURRENT APPLICATION NUMBER: 10/019,282
 13 <141> CURRENT FILING DATE: 2002-01-02
 15 <150> PRIOR APPLICATION NUMBER: PCT/JP00/04285
 16 <151> PRIOR FILING DATE: 2000-06-28
 18 <150> PRIOR APPLICATION NUMBER: JP189644/1999
 19 <151> PRIOR FILING DATE: 1999-07-02
 21 <160> NUMBER OF SEQ ID NOS: 76
 23 <170> SOFTWARE: PatentIn version 3.2
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 5692
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Streptomyces Sp.
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (948)..(3362)
 35 <400> SEQUENCE: 1
 36 gaattccgga tggttggaga ggccgatcca gacggtgggc ggggcgaaga ggctgtcggc     60
 38 caggcccgct tcgacgaggt cgaagatcga ggcggcgtcc ggaccgtcca ggatggtgtt    120
 40 ctccgcgccg accgccagat agggcagcag gaacacgtgc atctgggccg agtggtagag    180
 42 cggcagggag tgcacgggcc ggtcggtcgc ggcgaggccg agcgcggtga tcgcgctgac    240
 44 gtactcgtgg accagggccc cgtgcgtcat catcgcgccc ttgggcaggg cggtggtccc    300
 46 ggaggtgtac agcagctgca ccaggtcgtc ggaggcgggc gggcgccgcg gggtgaacgc    360
 48 ccgttccgtc tccagggcgt cgagcagcga gccgggcgcg tcgcggagcg cgcgcaccgg    420
 50 gagtccggcg gggagccgcc cggcgaggtc cgggtcggtc aggacgaggg aggagccgga    480
 52 ctggtcgagg aggtaggcca ggtcgtcgcc ggtgaggttc tggttgaccg gtacgtggac    540
 54 gagaccggcc cgtgcgcagg cgaggaagcc gatcagatag gcgtcggagt tgtgcgcgta    600
 56 ggcggccacc cggtcgccgg gggcgagagc gtactcctcg gtgaggacgg cggcggccgt    660
 58 ggagacggcg gcgtccaggg agcggtaggt ccaggtccgg tcggcgtacc gcacggcggt    720
 60 ccggtcgggg gtgcgccggg cgctgcgggt gaggacgccg tcgactgtgc tgctgcgtac    780
 62 acctgtcatg gcgtgatcct gtgcgtccgg gccctcgggg gtcaagaggc tggataccga    840
 64 ccagacggtt gacagcttcc cgggctccct ggctgagtga cgcttggccg tccgggcgtt    900
 66 ccggaccggc cgcgcccgtg ccacccgtac cgctgggagg aaacacc ttg acg tta     956
 67                                                    Leu Thr Leu
 68                                                    1
 70 cgc aac cgt ctg aga ctg ctc ggg gtc gcc ggt ctc gcc ctg ttc acc     1004
 71 Arg Asn Arg Leu Arg Leu Leu Gly Val Ala Gly Leu Ala Leu Phe Thr
 72     5                   10                  15
```

RAW SEQUENCE LISTING DATE: 10/19/2004
PATENT APPLICATION: US/10/019,282 TIME: 17:29:38

Input Set : D:\217770US0PCT.txt
Output Set: N:\CRF4\10192004\J019282.raw

```
74 gtg tcg gcg tcg ctg ccg cct gcc acc gcg tcc ggg acc cag gag acg     1052
75 Val Ser Ala Ser Leu Pro Pro Ala Thr Ala Ser Gly Thr Gln Glu Thr
76 20                  25                  30                  35
78 cgg cac ccg tcc ggg agc ggt ctt tcg gcc gtc atc cgg tac acg gag     1100
79 Arg His Pro Ser Gly Ser Gly Leu Ser Ala Val Ile Arg Tyr Thr Glu
80                  40                  45                  50
82 tac ggc att ccg cac atc gtg gcg gag gac tac gcg cag ttg ggc ttc     1148
83 Tyr Gly Ile Pro His Ile Val Ala Glu Asp Tyr Ala Gln Leu Gly Phe
84              55                  60                  65
86 ggc acc ggc tgg gcg cag gcc gcc gat cag gtg tgc acg ctg gcg gac     1196
87 Gly Thr Gly Trp Ala Gln Ala Ala Asp Gln Val Cys Thr Leu Ala Asp
88          70                  75                  80
90 ggc ttc ctc acg gtg cgc ggg gag cgg tcg agg ttc ttc ggc ccg gac     1244
91 Gly Phe Leu Thr Val Arg Gly Glu Arg Ser Arg Phe Phe Gly Pro Asp
92      85                  90                  95
94 gcc gcc acg gac tac tcc ctc tcc tcg gcg gcg acg aac ctc tcc agc     1292
95 Ala Ala Thr Asp Tyr Ser Leu Ser Ser Ala Ala Thr Asn Leu Ser Ser
96 100                 105                 110                 115
98 gac ctg tac ttc cgg ggc gtc cgc gac agc ggc acg gtg gag aag ctg     1340
99 Asp Leu Tyr Phe Arg Gly Val Arg Asp Ser Gly Thr Val Glu Lys Leu
100                 120                 125                 130
102 ctc aag gag ccc gcg ccc gcc ggt ccg agc agg gac gtc aag gag acg     1388
103 Leu Lys Glu Pro Ala Pro Ala Gly Pro Ser Arg Asp Val Lys Glu Thr
104             135                 140                 145
106 atg cgc ggg ttc gcc gcc ggg tac aac gcg tgg atc gcg cag aac cgg     1436
107 Met Arg Gly Phe Ala Ala Gly Tyr Asn Ala Trp Ile Ala Gln Asn Arg
108         150                 155                 160
110 atc acc gac ccc gcc tgc cgg ggc gcg tcc tgg gtg cgc ccg gtg acg     1484
111 Ile Thr Asp Pro Ala Cys Arg Gly Ala Ser Trp Val Arg Pro Val Thr
112     165                 170                 175
114 gcg ctg gac gtg gcg gcg cgc ggc tac gcg ctg gcg gtg ctc ggc ggc     1532
115 Ala Leu Asp Val Ala Ala Arg Gly Tyr Ala Leu Ala Val Leu Gly Gly
116 180                 185                 190                 195
118 cag ggg cgc ggc atc gac ggc atc acc gcg gca cag ccg ccg acc gcc     1580
119 Gln Gly Arg Gly Ile Asp Gly Ile Thr Ala Ala Gln Pro Pro Thr Ala
120                 200                 205                 210
122 gct cct ccg gcg gcc ggg gtc acg ccc gag gag gcg gcg acg gcg gcg     1628
123 Ala Pro Pro Ala Ala Gly Val Thr Pro Glu Glu Ala Ala Thr Ala Ala
124             215                 220                 225
126 gag cgg ctg ctg tcg acg cag aac gcg gac atg ggt tcc aac gcg gtg     1676
127 Glu Arg Leu Leu Ser Thr Gln Asn Ala Asp Met Gly Ser Asn Ala Val
128         230                 235                 240
130 gcc ttc gac ggc tcc acg acg gtg aac ggg cgc ggg ctg ttg ctc ggc     1724
131 Ala Phe Asp Gly Ser Thr Thr Val Asn Gly Arg Gly Leu Leu Leu Gly
132     245                 250                 255
134 aac ccg cac tac ccg tgg cag ggc gga cgc cgc ttc tgg cag gcg cag     1772
135 Asn Pro His Tyr Pro Trp Gln Gly Gly Arg Arg Phe Trp Gln Ala Gln
136 260                 265                 270                 275
138 cag acg atc ccc ggc gag ctg aac gtg tcg ggc gcg tcc ctg ctg ggc     1820
```

RAW SEQUENCE LISTING DATE: 10/19/2004
PATENT APPLICATION: **US/10/019,282** TIME: 17:29:38

Input Set : **D:\217770US0PCT.txt**
Output Set: **N:\CRF4\10192004\J019282.raw**

```
139 Gln Thr Ile Pro Gly Glu Leu Asn Val Ser Gly Ala Ser Leu Leu Gly
140                 280                 285                 290
142 gcg acg acg atc tcg atc ggg cac aac gcc gat gtg gcg tgg agc cat     1868
143 Ala Thr Thr Ile Ser Ile Gly His Asn Ala Asp Val Ala Trp Ser His
144             295                 300                 305
146 acg gtc gcc acg ggc gtc acg ctg aat ctg cat cag ctc agc ctc gat     1916
147 Thr Val Ala Thr Gly Val Thr Leu Asn Leu His Gln Leu Ser Leu Asp
148         310                 315                 320
150 ccg gcc gac ccg acc gtc tat ctg gtg gac ggg aag cgg gag cgg atg     1964
151 Pro Ala Asp Pro Thr Val Tyr Leu Val Asp Gly Lys Arg Glu Arg Met
152     325                 330                 335
154 acg cag cgg acg gtg agc gtc ccg gtg aag ggc ggg gcc gac gtg acc     2012
155 Thr Gln Arg Thr Val Ser Val Pro Val Lys Gly Gly Ala Asp Val Thr
156 340                 345                 350                 355
158 cgc acc cag tgg tgg acc cgc tac ggg ccg gtg gcc acc tcg atg ggc     2060
159 Arg Thr Gln Trp Trp Thr Arg Tyr Gly Pro Val Ala Thr Ser Met Gly
160                 360                 365                 370
162 gcg ggg ctg ccg ttg ccg tgg acg gcg agc acg gcg tac gcg ctg aac     2108
163 Ala Gly Leu Pro Leu Pro Trp Thr Ala Ser Thr Ala Tyr Ala Leu Asn
164             375                 380                 385
166 gat ccg aac gcg acg aat ctg cgg atg gcg gac acc ggt ctg ggc ttc     2156
167 Asp Pro Asn Ala Thr Asn Leu Arg Met Ala Asp Thr Gly Leu Gly Phe
168         390                 395                 400
170 ggc aag gcc cgc tcc acg ggt gac gtc gag cgt gcg ctg cac cgg tcg     2204
171 Gly Lys Ala Arg Ser Thr Gly Asp Val Glu Arg Ala Leu His Arg Ser
172     405                 410                 415
174 cag ggc atg ccg tgg gtg aac acg atc gcg gcg gac cgg gcg ggt cgc     2252
175 Gln Gly Met Pro Trp Val Asn Thr Ile Ala Ala Asp Arg Ala Gly Arg
176 420                 425                 430                 435
178 tcg ttc ttc gcg cag tcg cag gtg ctg ccg agg atc acc gac gcg ttg     2300
179 Ser Phe Phe Ala Gln Ser Gln Val Leu Pro Arg Ile Thr Asp Ala Leu
180                 440                 445                 450
182 gcg gag cgc tgc tcg acc ccg ctg ggc cgg gcc acc tac ccc gct tcc     2348
183 Ala Glu Arg Cys Ser Thr Pro Leu Gly Arg Ala Thr Tyr Pro Ala Ser
184             455                 460                 465
186 ggc ctc gcg gtg ctg gac ggt tcg cgg acg gac tgc gcg ctg ggc agc     2396
187 Gly Leu Ala Val Leu Asp Gly Ser Arg Thr Asp Cys Ala Leu Gly Ser
188         470                 475                 480
190 gac ccg gac gcg gtg cgg ccg ggg atc ttc ggc ccg ggc cgg atg ccg     2444
191 Asp Pro Asp Ala Val Arg Pro Gly Ile Phe Gly Pro Gly Arg Met Pro
192     485                 490                 495
194 gtg ctg aag aac cag ccg tac gtg gag aac tcc aac gac agc gcg tgg     2492
195 Val Leu Lys Asn Gln Pro Tyr Val Glu Asn Ser Asn Asp Ser Ala Trp
196 500                 505                 510                 515
198 ctg acc aat gcg gag cgg ccg ctg acc ggg tac gag cgg gtc ttc ggc     2540
199 Leu Thr Asn Ala Glu Arg Pro Leu Thr Gly Tyr Glu Arg Val Phe Gly
200                 520                 525                 530
202 acg atc gcg acg ccc cgg tcg atg cgg acg cgc ggc gcg atc gag gac     2588
203 Thr Ile Ala Thr Pro Arg Ser Met Arg Thr Arg Gly Ala Ile Glu Asp
```

RAW SEQUENCE LISTING DATE: 10/19/2004
PATENT APPLICATION: US/10/019,282 TIME: 17:29:38

Input Set : D:\217770US0PCT.txt
Output Set: N:\CRF4\10192004\J019282.raw

```
204             535                 540                 545
206 gtc gcg tcg atg gcg gac cgg ggc cgc ctc cgg gtc ggg gac ctt cag     2636
207 Val Ala Ser Met Ala Asp Arg Gly Arg Leu Arg Val Gly Asp Leu Gln
208         550                 555                 560
210 cgg cag cag ttc gcc aac cgt gcg ccg gcc ggg gat ctg gcc gcc tcc     2684
211 Arg Gln Gln Phe Ala Asn Arg Ala Pro Ala Gly Asp Leu Ala Ala Ser
212     565                 570                 575
214 gag gcc gcc aag tgg tgt gcg gcg ctg ccg ggc ggc acg gcc gtg ggc     2732
215 Glu Ala Ala Lys Trp Cys Ala Ala Leu Pro Gly Gly Thr Ala Val Gly
216 580                 585                 590                 595
218 tcc gac gga acg ccg gtc gac gtg tcg gcg gcc tgc cgg gtg ctg cgg     2780
219 Ser Asp Gly Thr Pro Val Asp Val Ser Ala Ala Cys Arg Val Leu Arg
220                 600                 605                 610
222 cgc tgg gac cgg acc gtg gac agc gac agc cgg ggc gcg ctg ctc ttc     2828
223 Arg Trp Asp Arg Thr Val Asp Ser Asp Ser Arg Gly Ala Leu Leu Phe
224             615                 620                 625
226 gac cgg ttc tgg cgg aag gcg tcg tcg gcg ccc gcc gcc gag ctg tgg     2876
227 Asp Arg Phe Trp Arg Lys Ala Ser Ser Ala Pro Ala Ala Glu Leu Trp
228         630                 635                 640
230 agg acg ccg ttc gat ccg gcc gac ccg gtg cgc act ccg cgc ggc ctg     2924
231 Arg Thr Pro Phe Asp Pro Ala Asp Pro Val Arg Thr Pro Arg Gly Leu
232     645                 650                 655
234 aac acg gcc gcg ccc gtc ctg ggc agg gcc ctg gcg gac gcc gtg gcg     2972
235 Asn Thr Ala Ala Pro Val Leu Gly Arg Ala Leu Ala Asp Ala Val Ala
236 660                 665                 670                 675
238 gag ctg cgg gcg gcg ggc atc gcg ctg gac gcc ccg ctg ggc gag cac     3020
239 Glu Leu Arg Ala Ala Gly Ile Ala Leu Asp Ala Pro Leu Gly Glu His
240                 680                 685                 690
242 cag ttc gtc gtg cgg aac ggc aag cgg ctc ccg atc ggc ggc ggg acg     3068
243 Gln Phe Val Val Arg Asn Gly Lys Arg Leu Pro Ile Gly Gly Gly Thr
244             695                 700                 705
246 gag tcg ctg ggc atc tgg aac aag acc gag ccg cag tgg aac gcg gcg     3116
247 Glu Ser Leu Gly Ile Trp Asn Lys Thr Glu Pro Gln Trp Asn Ala Ala
248         710                 715                 720
250 ggc ggc ggc tat acg gag gtg tcg tcg ggc tcc agc tac atc cag gcg     3164
251 Gly Gly Gly Tyr Thr Glu Val Ser Ser Gly Ser Ser Tyr Ile Gln Ala
252     725                 730                 735
254 gtc ggc tgg gac gac agc cgc tgc ccg gtg gcc cgg acg ctg ctg acg     3212
255 Val Gly Trp Asp Asp Ser Arg Cys Pro Val Ala Arg Thr Leu Leu Thr
256 740                 745                 750                 755
258 tac tcc cag tcg gag aac ccg aag tca ccg cac tac agc gac cag acc     3260
259 Tyr Ser Gln Ser Glu Asn Pro Lys Ser Pro His Tyr Ser Asp Gln Thr
260                 760                 765                 770
262 agg ctg tac gcg ggt gag cgc tgg gtg acg tcc cgg ttc tgc gag agg     3308
263 Arg Leu Tyr Ala Gly Glu Arg Trp Val Thr Ser Arg Phe Cys Glu Arg
264             775                 780                 785
266 gac atc gcg cgt tcg ccg gac ctg cgg gtg gtg cgg gtg cac gag cgg     3356
267 Asp Ile Ala Arg Ser Pro Asp Leu Arg Val Val Arg Val His Glu Arg
268         790                 795                 800
```

RAW SEQUENCE LISTING DATE: 10/19/2004
PATENT APPLICATION: US/10/019,282 TIME: 17:29:38

Input Set : D:\217770US0PCT.txt
Output Set: N:\CRF4\10192004\J019282.raw

```
270 cgg tag cgcggtgggc ggacgggccc gcccatccgc ggcgagaagg gcgtccgcct      3412
271 Arg
274 cggcgggcgc ccttctcacc gatgtgtcgt gaccgcgctc ccgggggcgt cctcaccgag   3472
276 ccgccgaagg gcccggcggc cgaacccgtg accatgcgtg cgacgcatca cgctccgtcg   3532
278 gctccgccct ccgcccgcgc ccaggccagc tgcgcgtcgc tcagcggcgg gtcgaagcct   3592
280 tccgggaaca gcagcatccg cggctgcggc cacatgttct ccggtccgtg ttcctgacag   3652
282 tccagggcga ggagatgcgg cccgtccccg caggactcgt gccggtaggg gcggtcgtgc   3712
284 gcccggcaga aatagccgaa caccgcacag tggtcgtcgc cgcccggtcg gtggaagccg   3772
286 gggtcgctga cgatcacggt caccggctcc tgccggttga gccgagggat gggccgggga   3832
288 tcacgccaca acagtcgagg agggagcaca cgctcatctt ccccggggcc gagcccacgg   3892
290 gaagggggag cacggcggga cgcctcccgt cggcgtgatc gaccgggccg tcccgctcgc   3952
292 gggcgggccc tcccggaccc gttgctctac agcgggcgct cgaagccctc ccagtacggt   4012
294 tcgcgcagcc gccgtttgta gagcttgccg ttggggtcgc ggggcatggc ggtgatgaag   4072
296 tcgaggctcc ggggtcgttt gtagccggcg agccgctcct cgcagtgggc gaggatcgcg   4132
298 gcggcgagcg cgggtgacgg ctcgtggcca tcggccggtt cgacgacggc cttgacctcc   4192
300 tcgccgcggt cggcgtgggg gatgccgaag gcggcggcgt ccgcgacggc ggggtgggtg   4252
302 agcaggaccg actcgatctc ggcggggtag atgttgaccc cgcccgcgat gatcatgtcg   4312
304 atcttgcggt cgcggaggaa gaggtagccg tcctcgtcca gcacgccgag gtcaccgacg   4372
306 gtgaagaagt cgccgatgcg gttcgtgcgg gtcttggtct cgtccttgtg gtagctgaag   4432
308 ccgccggtgc tcatcttcat gtagacggtg cccagttcgc ctggcgggag gcggttgccg   4492
310 tcgtcgtcga agacggccag ttcgctgatc ggccaggcct tgccgacggt gccgggcttc   4552
312 ttcagccagt cctcggcggt ggcgaacgct cccccgccct cgctggccgc gtagtactcc   4612
314 tcgacgcagc tcccccacca gtcgatcatg gcgcgtttga cgtggtcggg gcaggggggct   4672
316 gccccgtgga tggcgtgccg catggaggag acgtcgtagc gggacctcac ctcgtcgggc   4732
318 agcgcgagca gccggtggaa ctgggtgggg accatgtggg tgtgggtgca gcggtgggcg   4792
320 tcgacgaggc gcagcatctc ctcgggcgac cagccgtcca tcaggaccag cgggtggccg   4852
322 atgtgcaggg cggcgcccgc gaattggagt acggcggtgt ggtagagcgg cgagcagacc   4912
324 aggtggacgt tgtcgtcgaa cggccggatg ccgaagatgc cgaggaaccc gccgaggtag   4972
326 gtctcctcgg ggcgtttgcc gggcaggggg cgccggatgc cgcgggggcg gccggtggtg   5032
328 cccgaggtgt agttcatgac ccagccgagg gtgcggttct caggcggcgt ggcggggtgg   5092
330 ccttcgagga gttcggccca ggggcggcag ccggggaccg tgccgacgcc gtagcggtgg   5152
332 gtcgcgggca gttccgcctc gtcggcggcg gccgtcgcgg tggccgcgaa gcgttcgtgg   5212
334 gcgatcagga cgcgggcgcc ggagtcggcg acgatccagg cgatctcggg gccgacgagg   5272
336 tggtggttga ccggcacgag gtagaagccg gcctgcgagg cggcgaggtg ggcggtgagg   5332
338 agttcgacgc cgttgggcag gacgacggcg aacgcgtcgc cctcgcgcag tccggccgcg   5392
340 cgcaggccgt ggaccatgcg gttgacgtcg gcgtgcaggc ggcccgcgct ccactcctcg   5452
342 ccgtcggggg cgatcaggac ggtgcggtcg gggtcggctg cggcctgggc ccagaaaccg   5512
344 ttgggcggct ggttcacgtg gcactccttc cggcgatgcg gttcatgcgg gtgacggccc   5572
346 gttcgaagcc gcgggtcagg tcgtcgacga cggcccggac gctgcgttca ctggtcatcc   5632
348 ggccgacgat ctgcccgacg ggcgtgccga gcagctcgcc gacctcgtac ttctggatcc   5692
351 <210> SEQ ID NO: 2
352 <211> LENGTH: 804
353 <212> TYPE: PRT
354 <213> ORGANISM: Streptomyces Sp.
356 <400> SEQUENCE: 2
358 Leu Thr Leu Arg Asn Arg Leu Arg Leu Leu Gly Val Ala Gly Leu Ala
359 1               5                  10                  15
362 Leu Phe Thr Val Ser Ala Ser Leu Pro Pro Ala Thr Ala Ser Gly Thr
363             20                  25                  30
```

VERIFICATION SUMMARY DATE: 10/19/2004
PATENT APPLICATION: **US/10/019,282** TIME: 17:29:39

Input Set : **D:\217770US0PCT.txt**
Output Set: **N:\CRF4\10192004\J019282.raw**